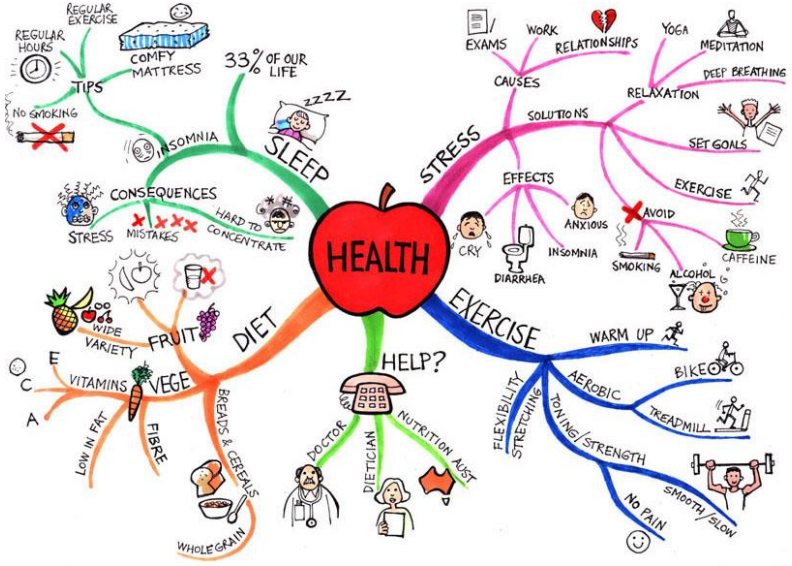
Reviewing GNNs' Developments from 2016 to 2023

Yiwei WANG
April 2024

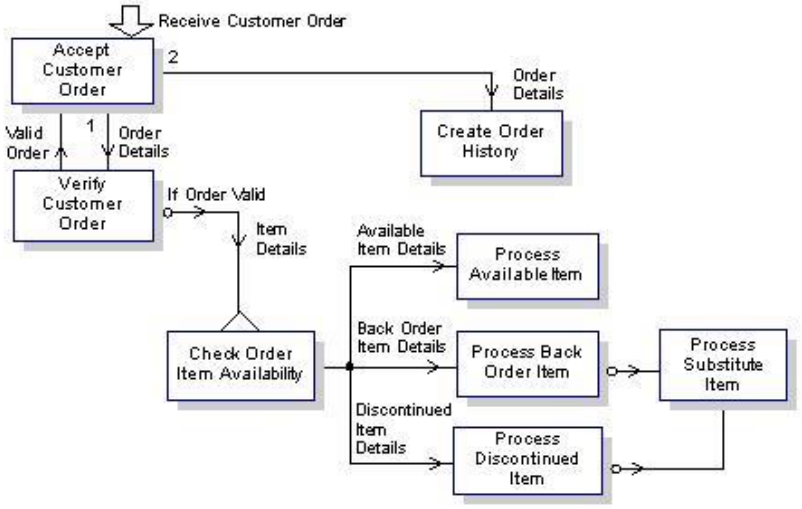
Human Intelligence with Relation Understanding

Different concepts in the real world are not isolated but connected. The understanding of relations help to build humans' intelligence. Question:

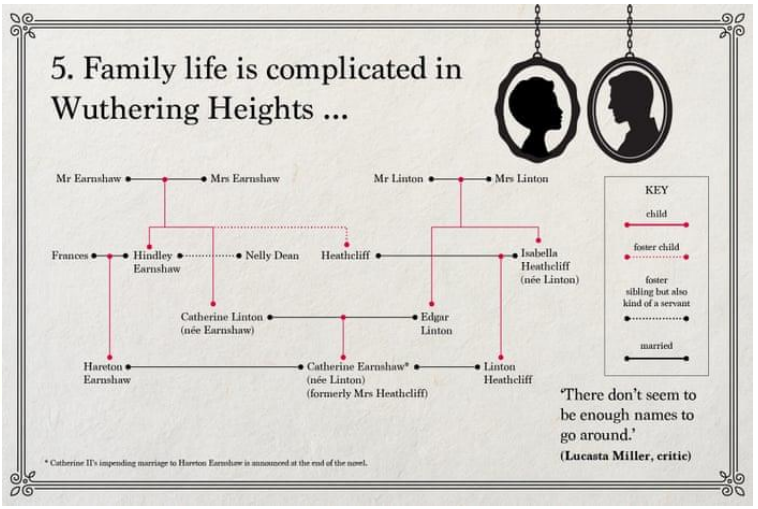
Can we also enable Artificial Intelligence to learn from some data to understand the relations of concepts?



A mindmap graph of healthcare



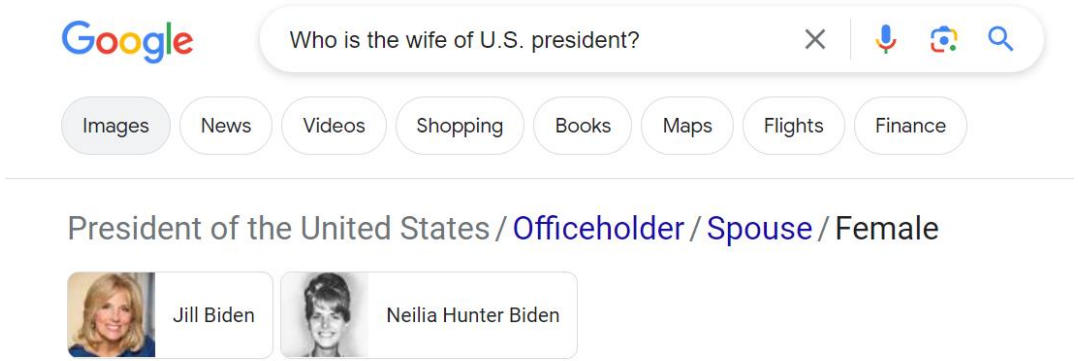
A software dependency graph



A people relation graph

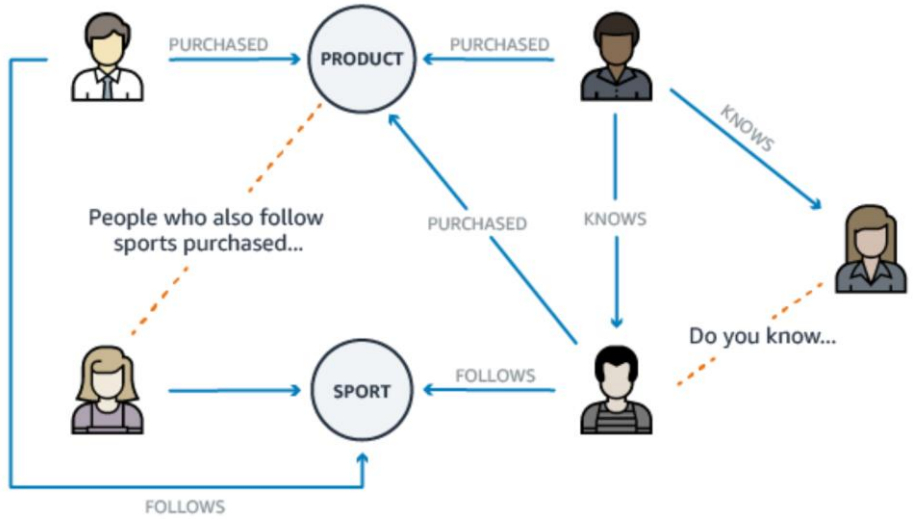
AI Also Needs to Understand Relations of Concepts

QA & Semantic Search



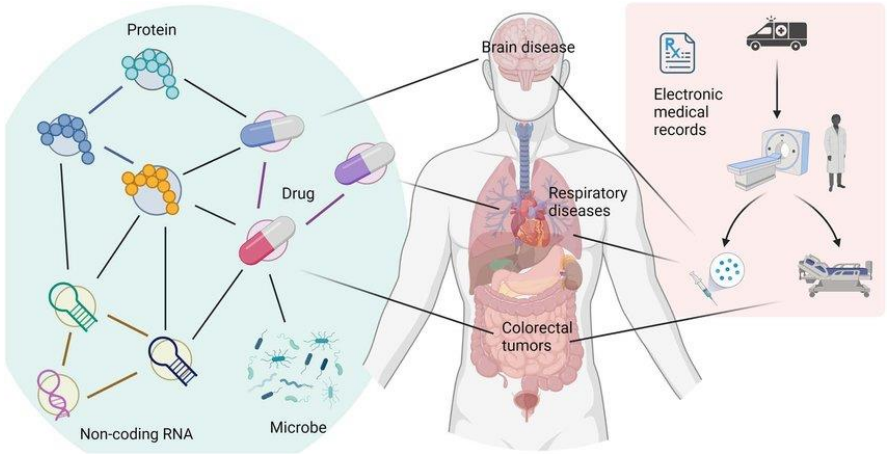
Relations between Entities

E-Commerce



Relations between Buyers, Sellers, Products

Computational Biomedicine

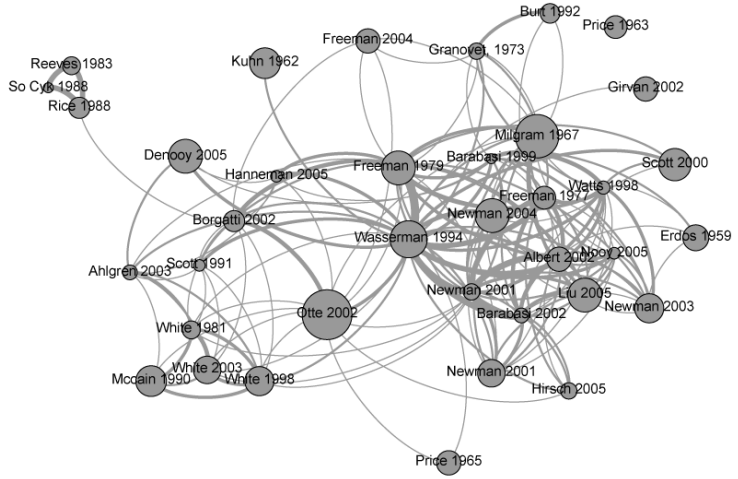


Interactions of (bio)molecules

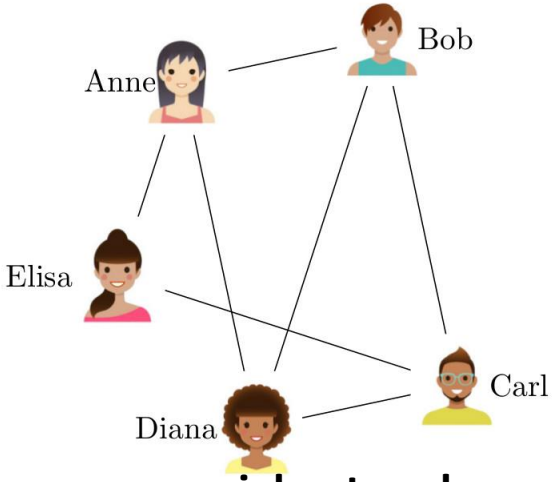
Relations of diseases and drugs

Graphs Represent Relations of Concepts

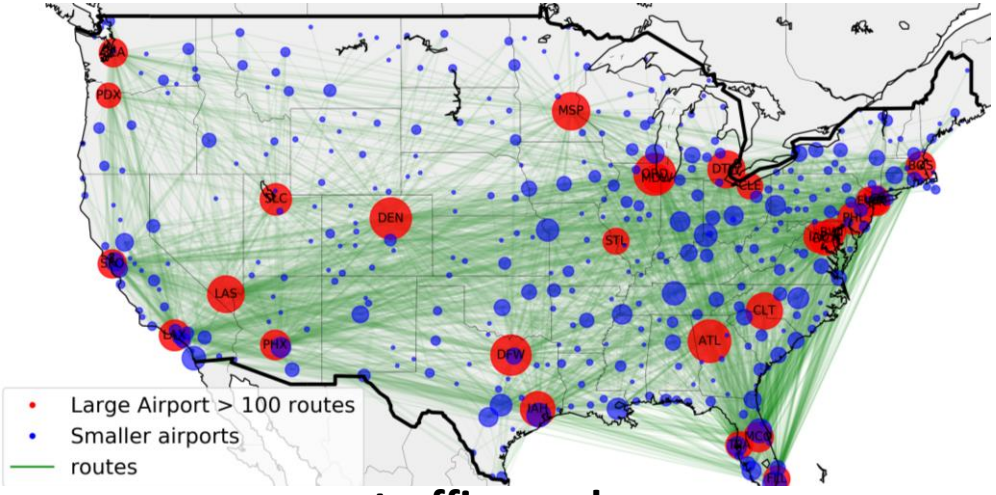
A typical way to represent relations between concept is using a graph, where a node is a concept and an edge represents a relation. This unified representation can be used to describe …



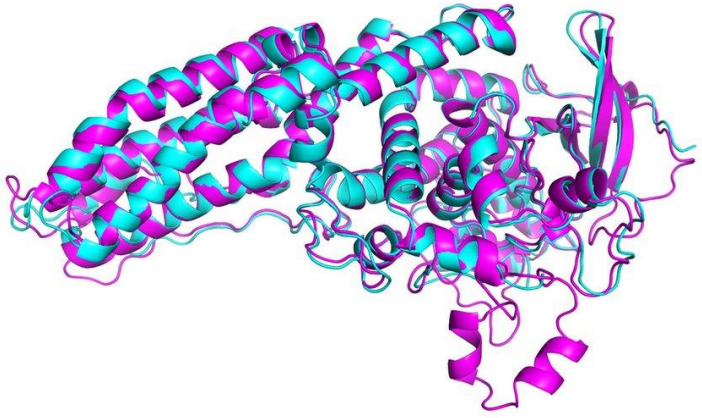
citation graph



social network



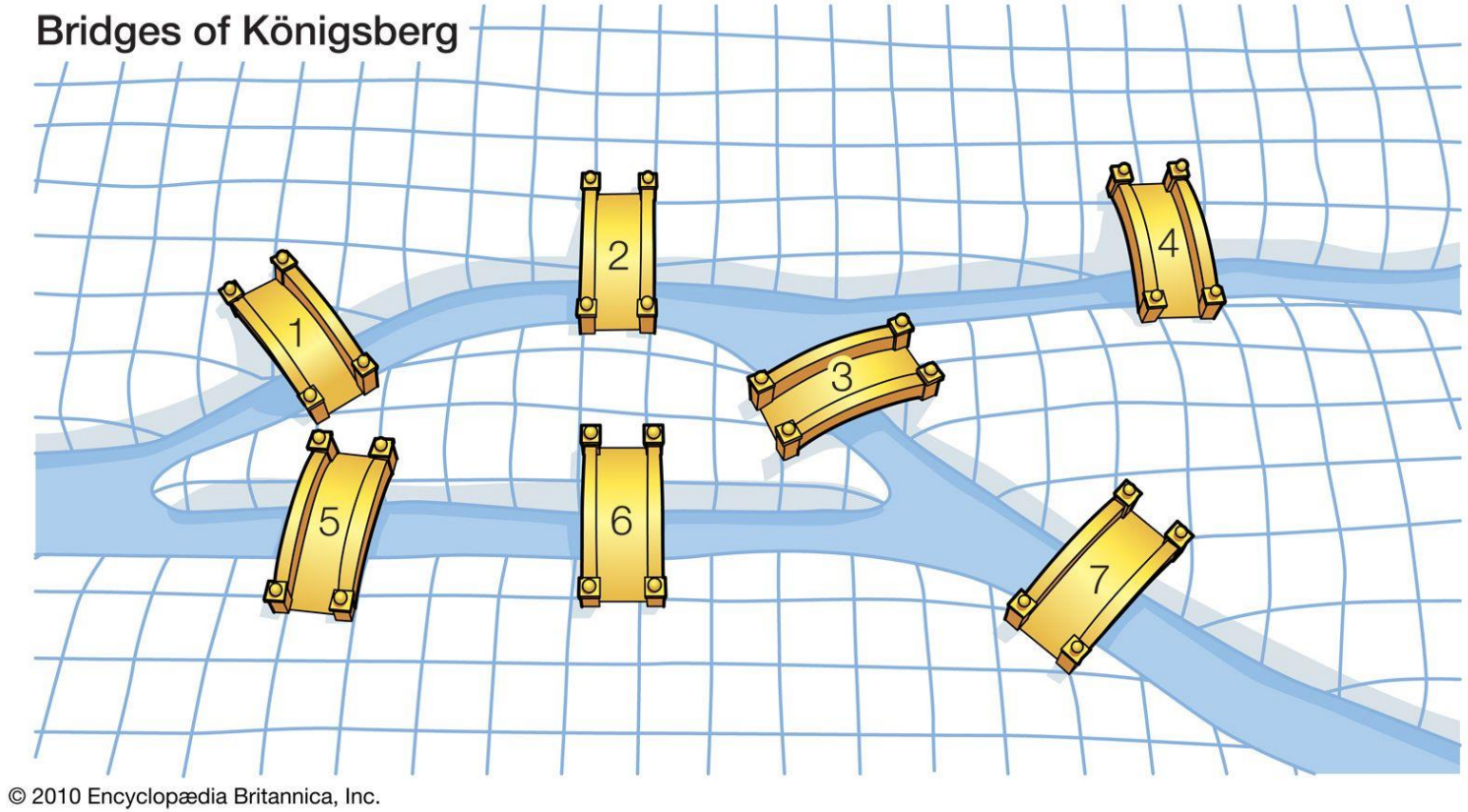
traffic graph



biology graph

Some Fundamental Questions to Ask to Know Graphs Better

What is the earliest graph in humans' history?



The earliest known example of a problem that can be considered a graph problem dates back to **1736**, when the Swiss mathematician Leonhard Euler solved the famous "Seven Bridges of Königsberg" problem.

https://www.britannica.com/topic/graph-theory

Graphs are Young and Advanced

Vision data exist before the human being's existence. Language data arises at the beginning of human being's wisdom.

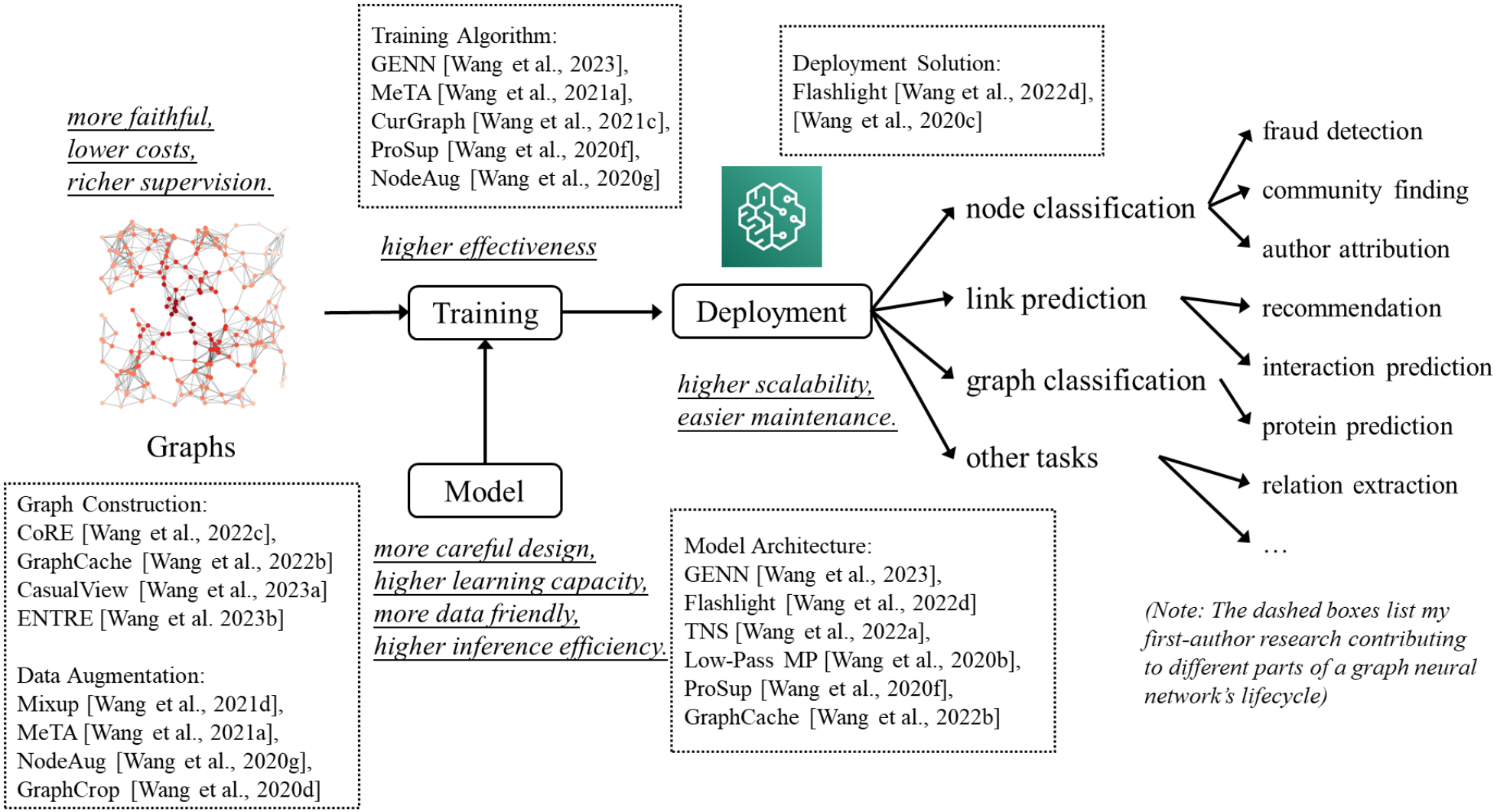
Different from the above data, graph data is an advanced kind of data that humans propose in the development of mathematics to describe the connections between concepts in the real world. This property determines the developments of GNN:

1. The graphs' capability of describing connections is the basis of GNNs' high performance on semi-supervised learning.

2. The ''expert-made'' property makes graph data limited in nature.

Lifecycle of a GNN Model

The lifecycle of a graph machine learning model generally includes following stages: graph data preparation, model design, model training, and model deployments.



Challenges of Doing Graph Machine Learning

1. Limited Graph Data
2. Expensive Supervision
3. Increasing Requirements of Privacy Protection and Low Inference Latency
4. High Complexity and Irregularity of Graphs Induces More Difficulty

Challenge 1: Limited Graph Data

Much data is not born to be structured. Most real-world data is unstructured text. To enable AI to understand the relations of concepts in the massive unstructured text, we have to first extract the relations from the text.

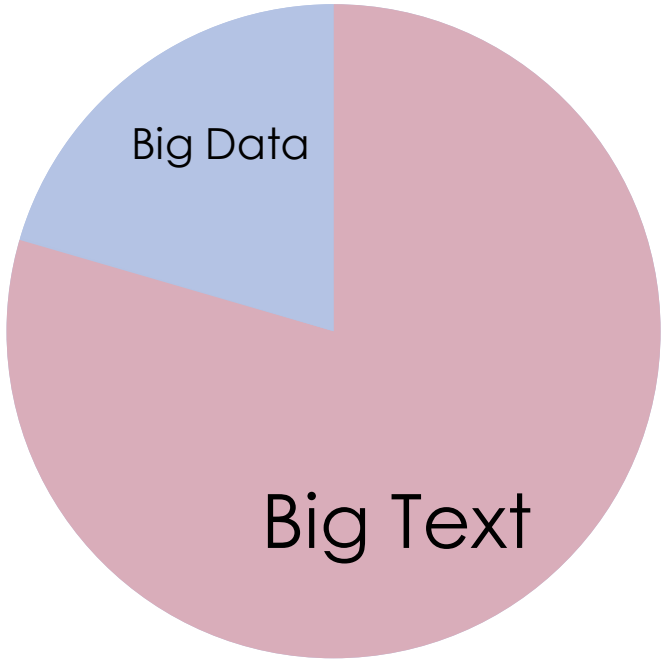
Every single day

1200 news stories

4000 papers

2M blog posts

500M tweets

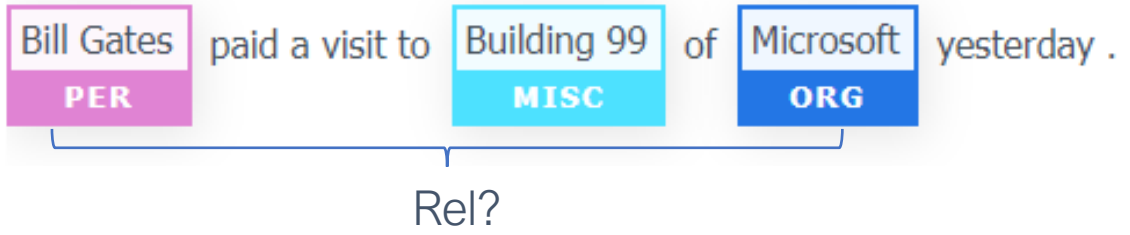


>80% of the world's data unstructured text
[by breakthrough analysis]

Faithfulness Issues of Graph Construction

Relation extraction aims to identify the relations described by the input text. However, relation extraction may not extract what is described in the text.

Relation extraction:



Visit ✓

FounderOf ✗

According to prior knowledge

Prior knowledge in large language models may lead to unfaithful extraction.

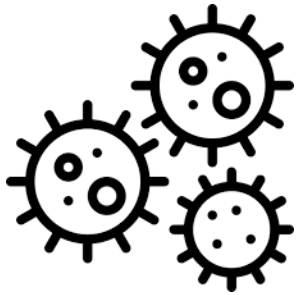
Context: Bill Gates went to Microsoft Building 99.
Question: What's the relation between Bill Gates and Microsoft in the given context?
Option: founder, visitor.
Answer with one word: founder (GPT-3.5) ❌

Why is Faithful Graph Construction so Important?

Drug-drug Interaction

The amount of metformin absorbed while taking Acarbose was bioequivalent to the amount absorbed when taking placebo, as indicated by the plasma AUC values. However, the peak plasma level of metformin was reduced by approximately 20% when taking Acarbose due to a slight delay in the absorption of metformin.

Relation type: mechanism



Disease-target detection

TOMM70, the most frequent binding partner of SARS-CoV-2 ORF9b, was identified in more than 1000 PSMs of the prey.

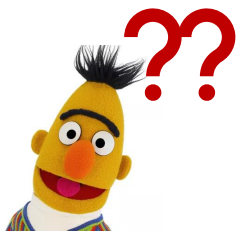
Relation type: binding

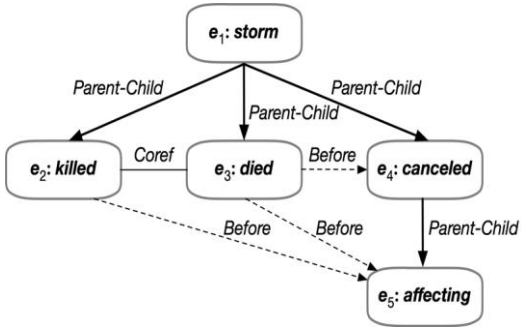
More risky tasks where we couldn't afford any ***GUESSES*** of **unfaithful Graph Construction**

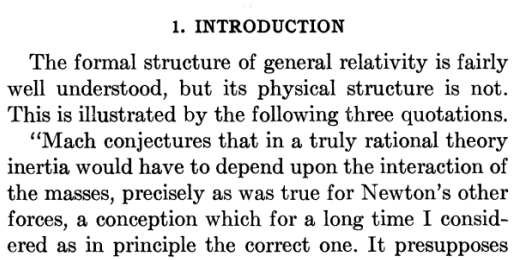
- **Disease phenotype extraction** from medical reports
- **Disaster and location extraction** from social media
- **API version compatibility** detection from software documents
- **Travel departure and arrival extraction** from emails and meeting logs
- …

Challenge 2: Expensive Supervision

The annotations of graph data is time-consuming, expensive, and usually require expert knowledge.



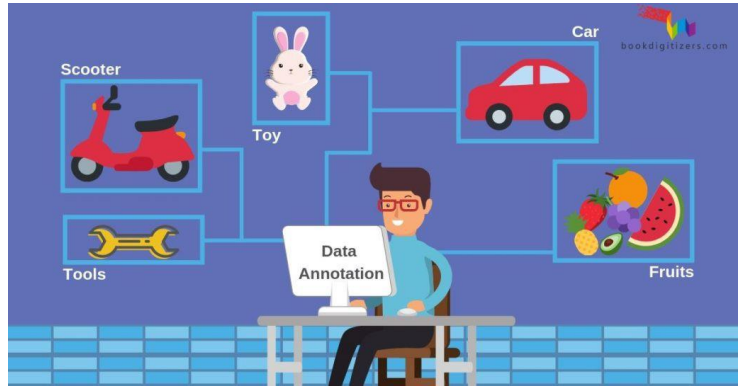




1. INTRODUCTION

The formal structure of general relativity is fairly well understood, but its physical structure is not. This is illustrated by the following three quotations.

"Mach conjectures that in a truly rational theory inertia would have to depend upon the interaction of the masses, precisely as was true for Newton's other forces, a conception which for a long time I considered as in principle the correct one. It presupposes



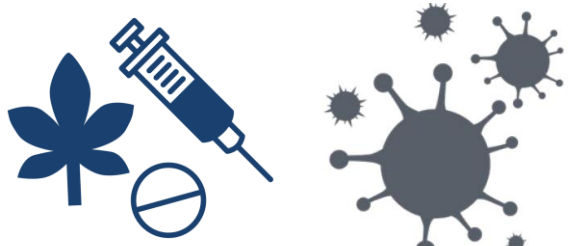
- Reading long text, history log, recognizing complex structures

>10 minutes of multiple expert investigators for just 1 label for node classification in Amazon E-commerce

Insufficiency

- **General domain**: 70million nodes and 18 billion edges in Wikidata.
- . **Specialized domains**: Millions of sellers and billions of edges in Amazon E-commerce graphs。

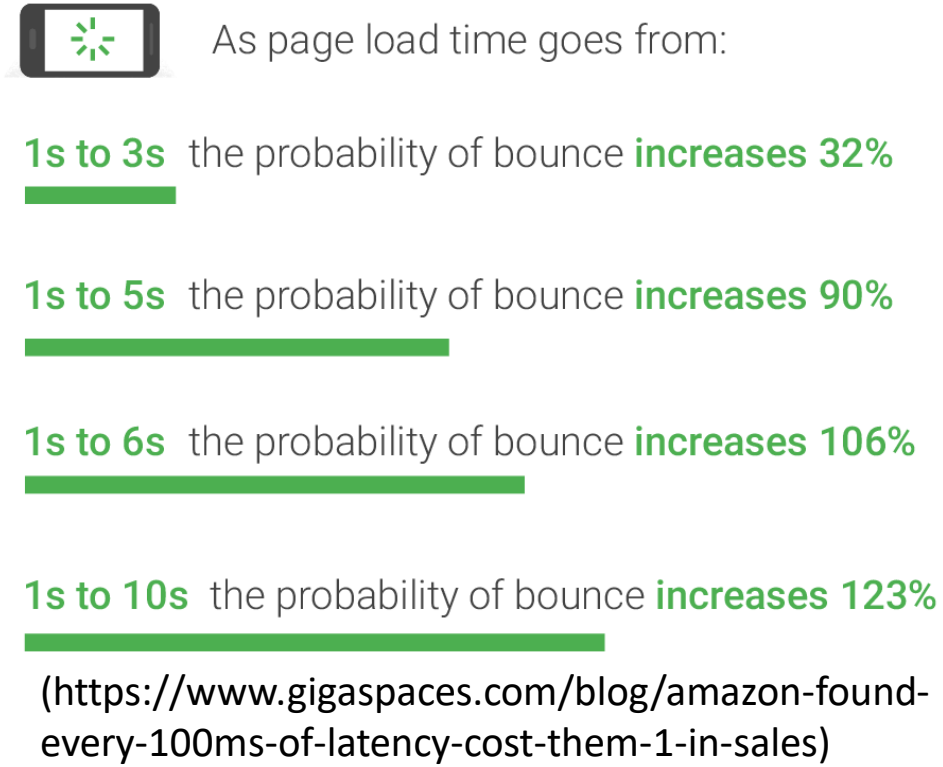
Low-resource Domains with Limited Annotations



Challenge 3: Increasing Requirements of Privacy Protection and Low Inference Latency

Consumers and companies desire more privacy protection and lower inference latency of AI services.

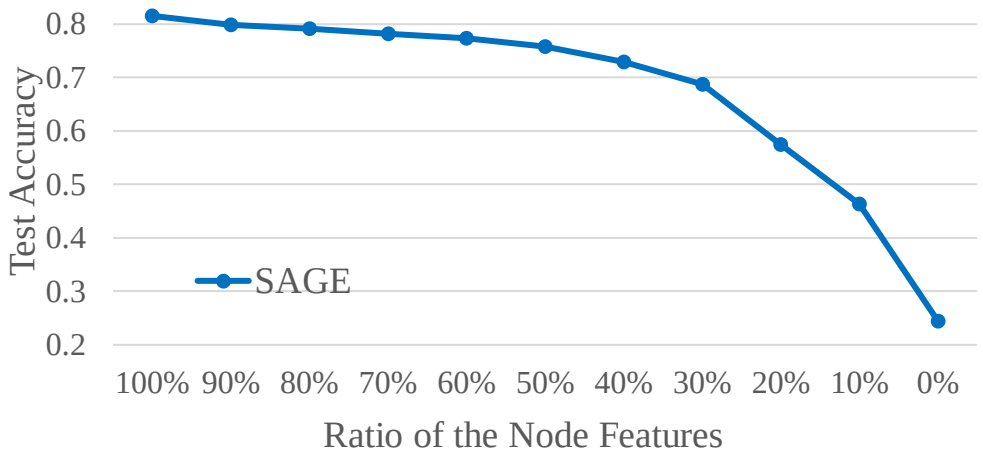
(https://www.technewsworld.com/story/apple-privacy-rule-cost-tech-titans-estimated-9-85-billion-in-revenue-87324.html)



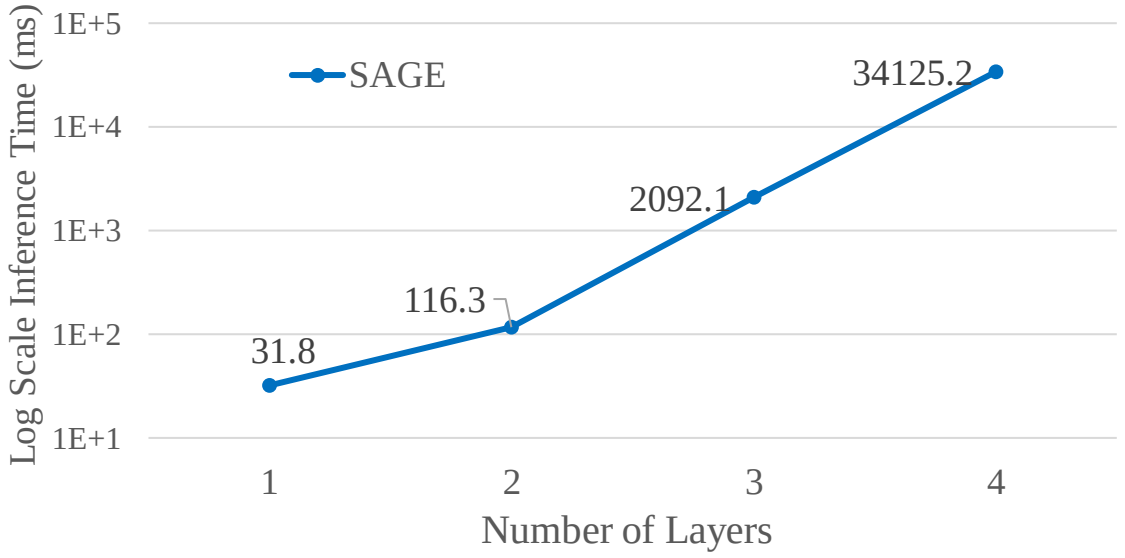
(https://www.gigaspaces.com/blog/amazon-found-every-100ms-of-latency-cost-them-1-in-sales)

Prior Graph Neural Networks Struggle for Privacy Protection and Low Inference Latency

Serious accuracy degradation with fewer node features.

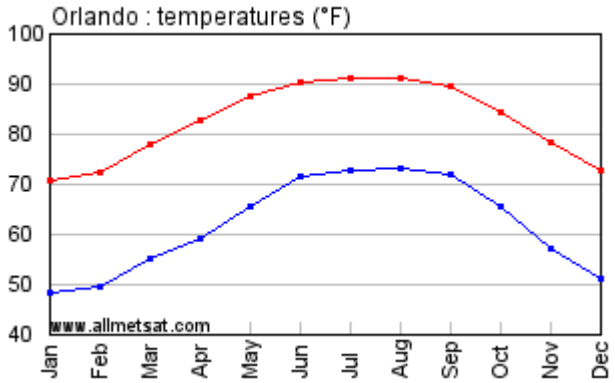


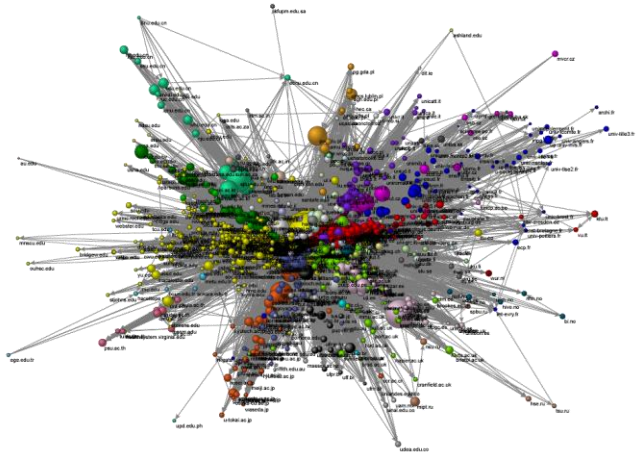
Inference latency grows exponentially as the number of layers increases.

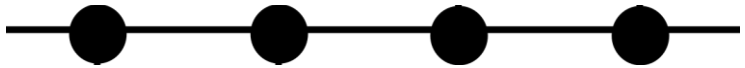


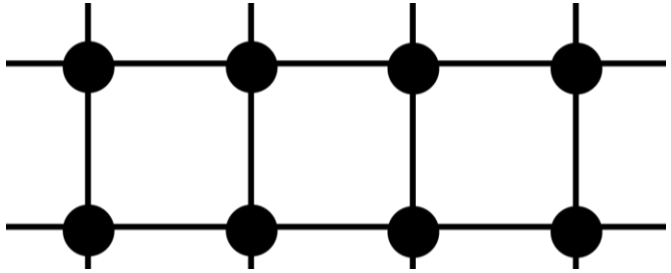
Challenge 4: High Complexity and Irregularity of Graphs Induces More Difficulty

Time series data is on a one-dimensional grid. Images are on a two-dimensional grid. Different from them, graph data cannot be placed on a regular grid. The high complexity and irregularity of graph data makes the design of machine learning algorithms much more challenging.









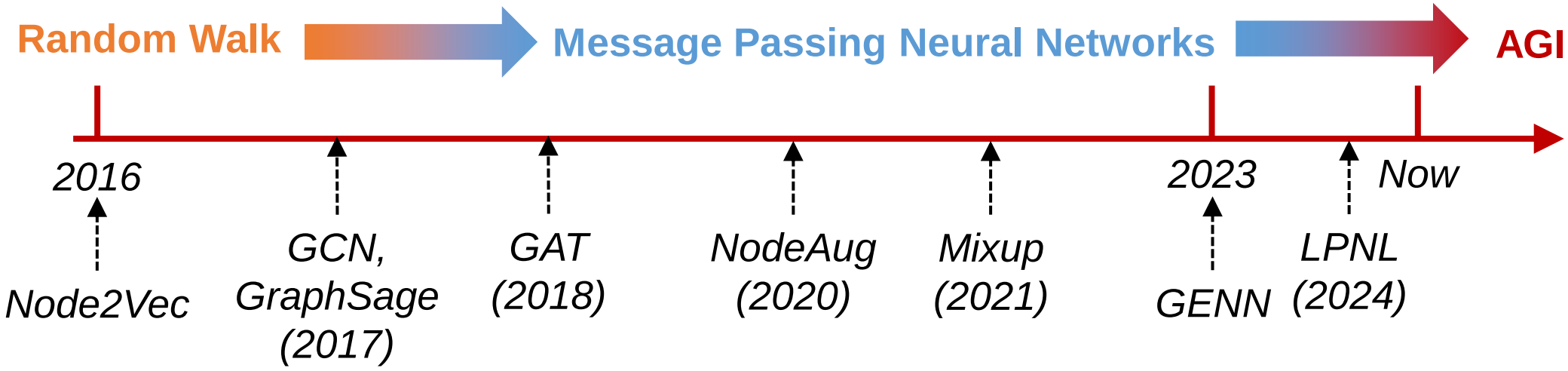
Reviewing GNNs' History to Look Ahead

"Taking history as a mirror, we can understand prosperity and decline."
- Zheng Wei (580 A.D. – 643 A.D.)

My View of Fundamental Paradigms' Evolution in GNNs:

From 2016 to 2023, From Random Walk *to* Message Passing Neural Networks
Influence: Improved Learning Capability on Attributed Graphs.

From 2023 to ***future****, From* Message Passing Neural Networks *to* **Artificial General Intelligence (AGI)**
Influence 1: Instruction-Tuned LLMs Provide A Unified Protocol to Learn from All Data Including Graphs.
Influence 2: From Explicit Graphs to Implicit Graphs.



node2vec: Scalable Feature Learning for Networks (2016)

node2vec: Scalable Feature Learning for Networks

Aditya Grover
Stanford University
adityag@cs.stanford.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

1. Node2vec borrows the idea from word2vec in NLP. It takes the nodes as the tokens and the random walk paths as the sentences to train the node-level embeddings.

2. Node2vec is an important baseline on learning from unattributed graphs.

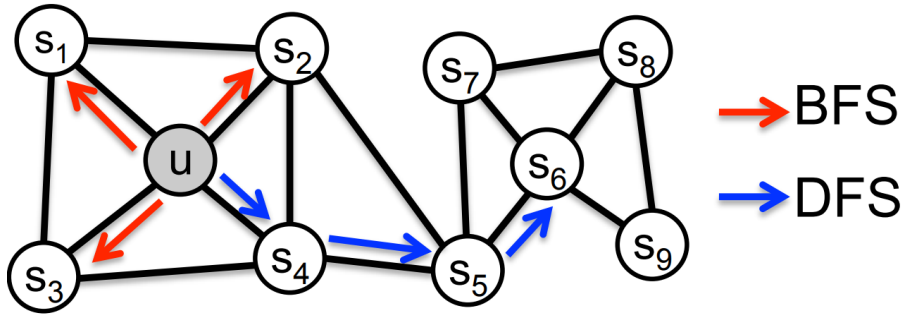


Figure 1: BFS and DFS search strategies from node u ($k = 3$).

SEMI-SUPERVISED CLASSIFICATION WITH GRAPH CONVOLUTIONAL NETWORKS (2017)

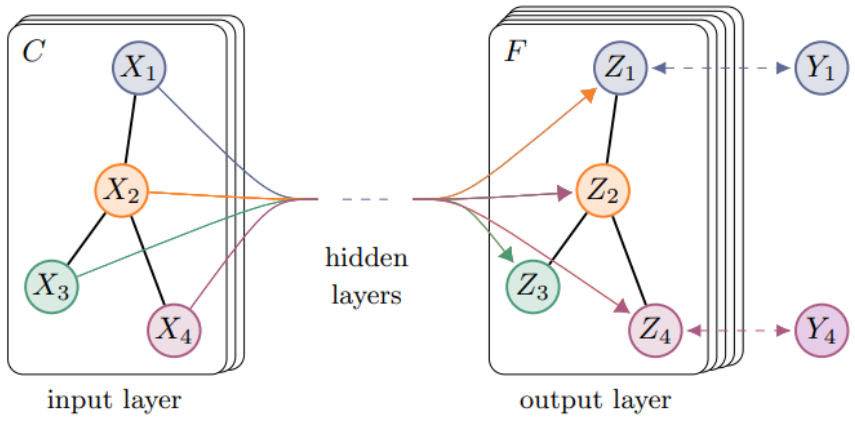
SEMI-SUPERVISED CLASSIFICATION WITH GRAPH CONVOLUTIONAL NETWORKS

Thomas N. Kipf
University of Amsterdam
T.N.Kipf@uva.nl

Max Welling
University of Amsterdam
Canadian Institute for Advanced Research (CIFAR)
M.Welling@uva.nl

1. GCN should be seen as the first graph neural network that is widely recognized by the community and starts the era of neural networks on graphs.

2. GCN, at the first time, shows the power of semi-supervised learning with graphs and neural networks.



(a) Graph Convolutional Network

GraphSage: Inductive Representation Learning on Large Graphs (2017)

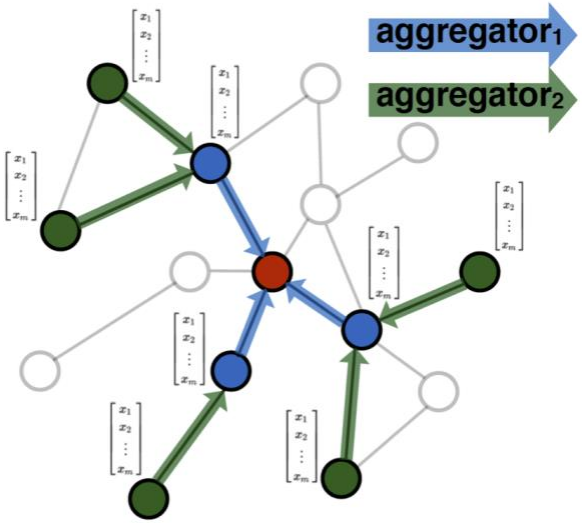
Inductive Representation Learning on Large Graphs

William L. Hamilton* wleif@stanford.edu
Rex Ying* rexying@stanford.edu
Jure Leskovec jure@cs.stanford.edu

Department of Computer Science
Stanford University
Stanford, CA, 94305

1. GraphSage is the first paper that defines the ‘’message passing” paradigm for the community, which is the basis of most GNNs.

2. GraphSage is the first inductive GNN.



GAT: GRAPH ATTENTION NETWORKS (2018)

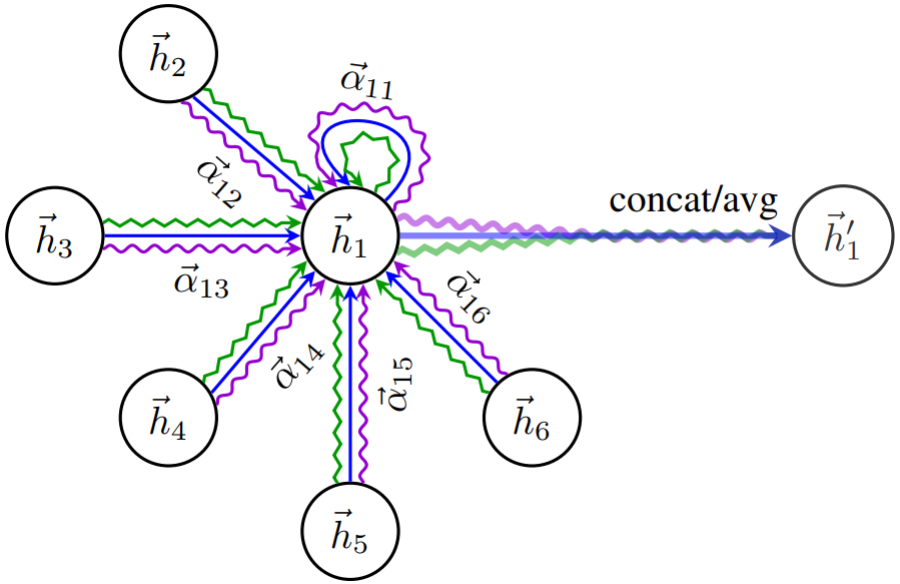
GRAPH ATTENTION NETWORKS

Petar Veličković*
Department of Computer Science and Technology
University of Cambridge
petar.velickovic@cst.cam.ac.uk

Guillem Cucurull*
Centre de Visió per Computador, UAB
gcucurull@gmail.com

1. GAT changes the message aggregator from the mean function to an attention function.

2. GAT explores the influence of edge weights in the message passing.



NodeAug: Semi-Supervised Node Classification with Data Augmentation (2020)

NodeAug: Semi-Supervised Node Classification with Data Augmentation

Yiwei Wang
National University of Singapore
Singapore
wangyw_seu@foxmail.com

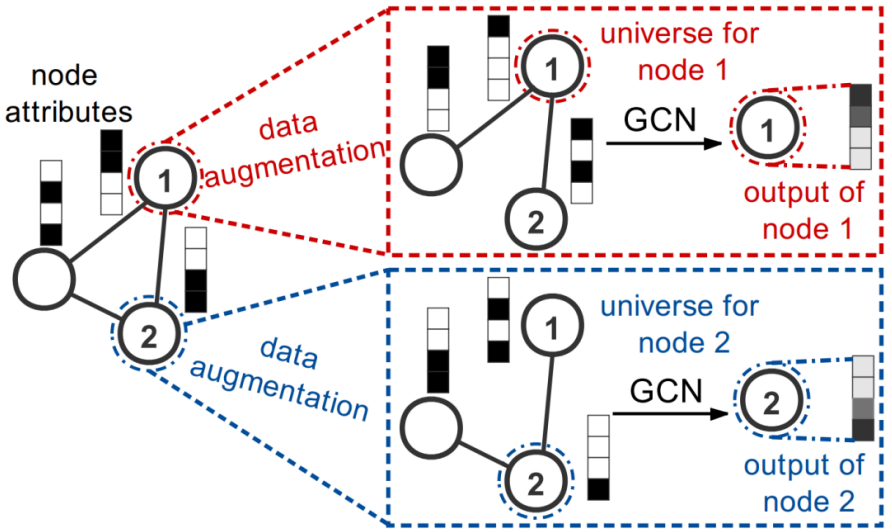
Wei Wang
National University of Singapore
Singapore
wangwei@comp.nus.edu.sg

Yuxuan Liang
National University of Singapore
Singapore
yuxliang@outlook.com

1. NodeAug is the first data augmentation framework for semi-supervised learning of GNNs.

2. NodeAug, at the first time, defines the receptive fields of GNNs.

3. NodeAug boosts the performance of GNNs on semi-supervised learning with data augmentation, motivates many data augmentation research on GNNs.



Mixup for Node and Graph Classification (2021)

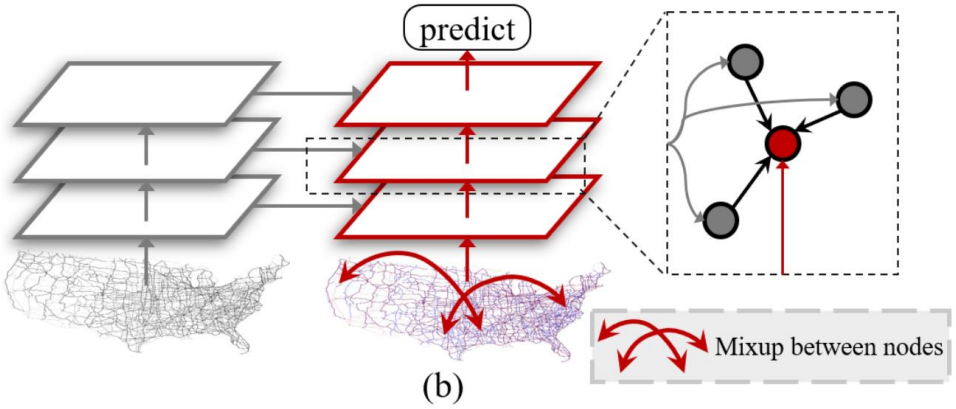
Mixup for Node and Graph Classification

Yiwei Wang
National University of Singapore
Singapore
wangyw_seu@foxmail.com

Wei Wang
National University of Singapore
Singapore
wangwei@comp.nus.edu.sg

Yuxuan Liang
National University of Singapore
Singapore
yuxliang@outlook.com

1. Mixup, at the first time, proposes the Mixup data augmentation method for GNNs.
2. Mixup is the first risk-free data augmentation method for GNNs.
3. Mixup boosts the performance of wide GNNs with negligible training costs.



THE WEB CONFERENCE

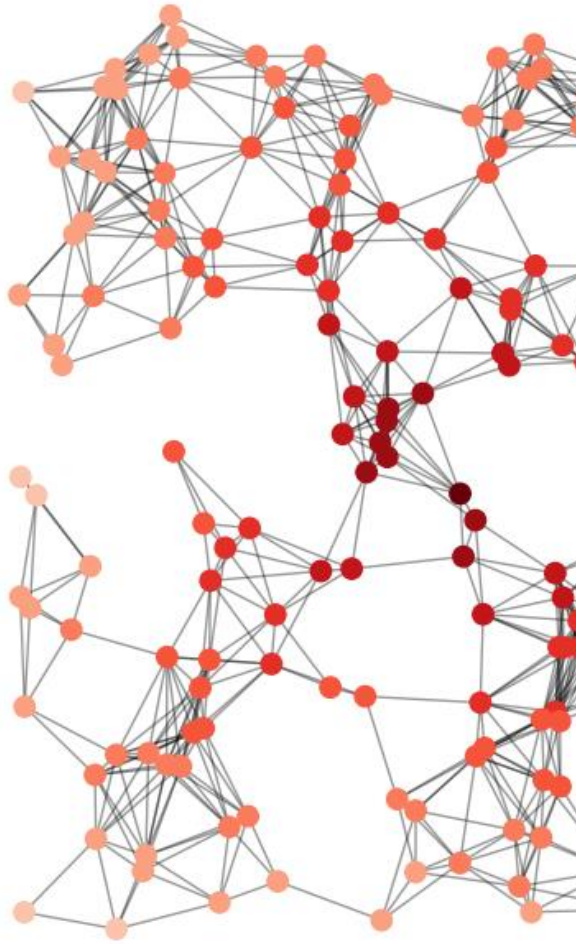
Mixup for Node and Graph Classification

The first Mixup data augmentation method for graph learning.

Graph Data Augmentation

Graph data augmentation improves the performance of graph learning with

- fewer human annotations
- lower costs
- richer supervision



What is Mixup?

Mixup is an advanced data augmentation technique originally proposed for Image Classification. In effect, Mixup regularizes the image classification models to favor simple linear behavior in between training examples and improves generalization.

$$\tilde{x} = \lambda x_i + (1 - \lambda) x_j, \qquad \text{where } x_i, x_j \text{ are raw input vectors}$$
$$\tilde{y} = \lambda y_i + (1 - \lambda) y_j, \qquad \text{where } y_i, y_j \text{ are one-hot label encodings}$$

Cohen, R., Biran, E., Yoran, O., Globerson, A., & Geva, M. (2023). Evaluating the Ripple Effects of Knowledge Editing in Language Models. arXiv preprint arXiv:2307.12976.

People Desire a Mixup Method for Graphs

Traditional Data Augmentation v.s. Mixup

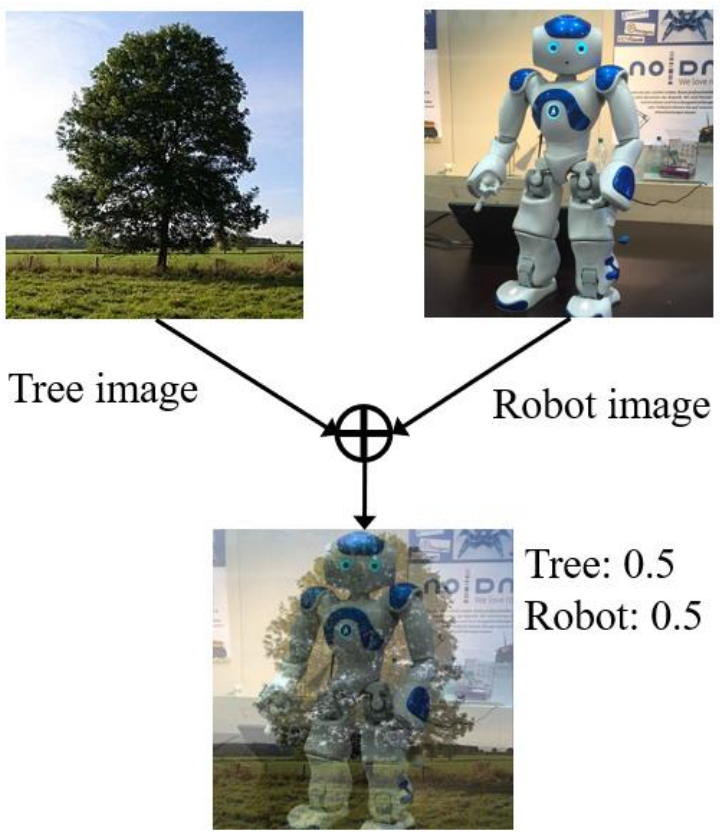
Risk of Breaking Ground-truth Labels v.s. Risk Free
Single Instance v.s. Augmentation Across Instances
Requiring Expert Knowledge v.s. No Requirements
Cannot Fit All Scenarios v.s. Scenario Agnostic
…

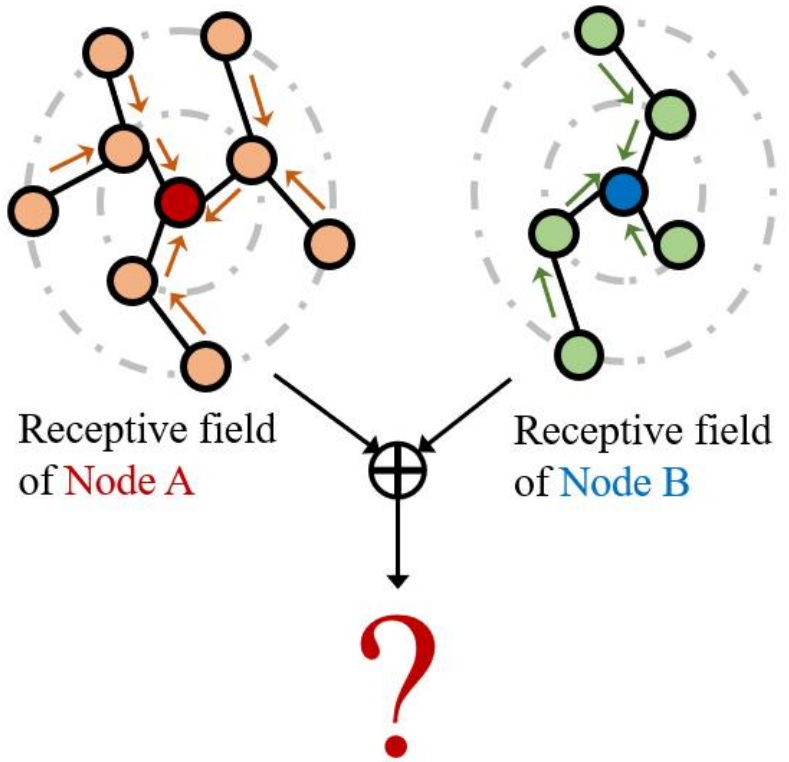
Overall, Mixup is more advanced, and people desire a Mixup method for Graphs!

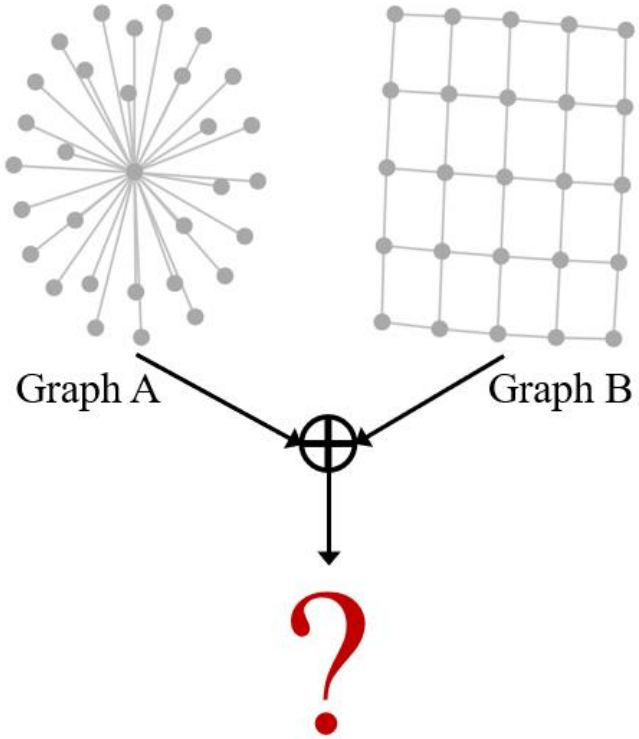
But …

Designing Mixup methods is very Challenging for Graphs

The challenges are rooted in the irregularity and connectivity of graph data.

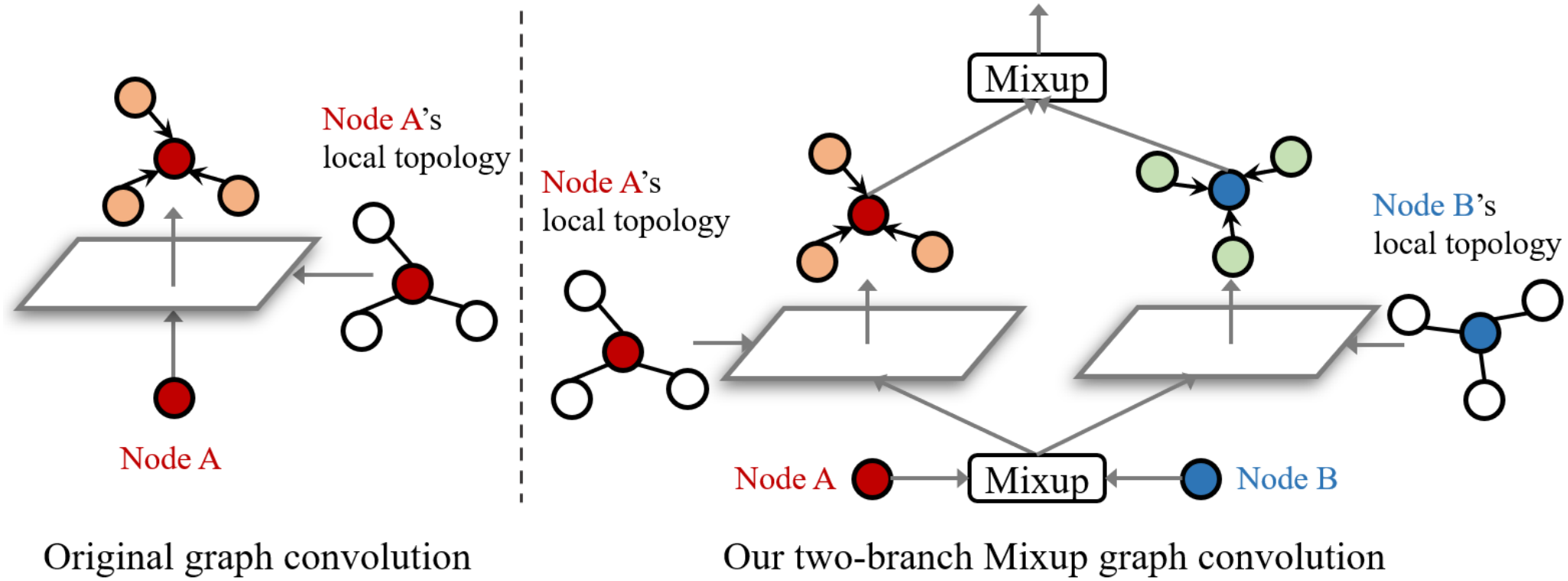




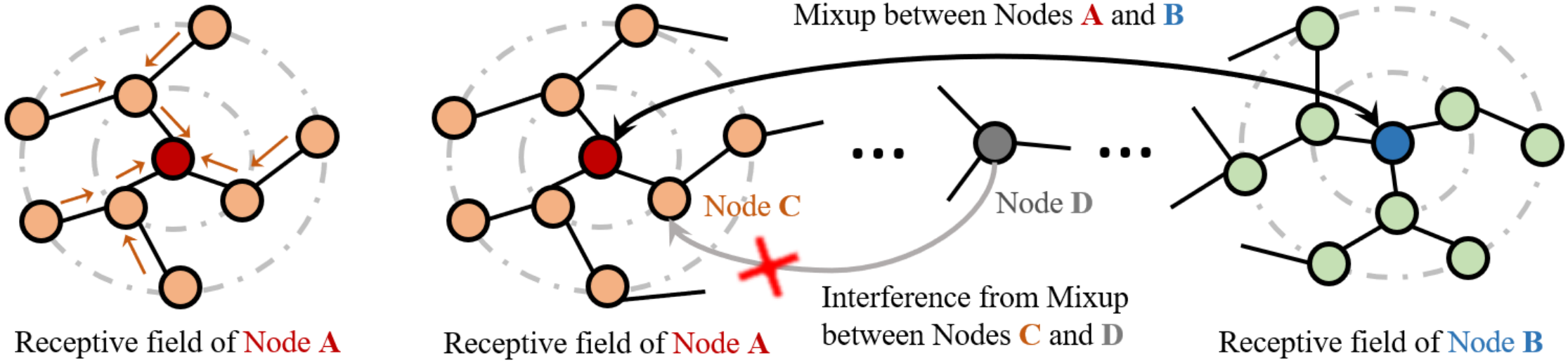


Our Work: The First Mixup Method for Graph Learning

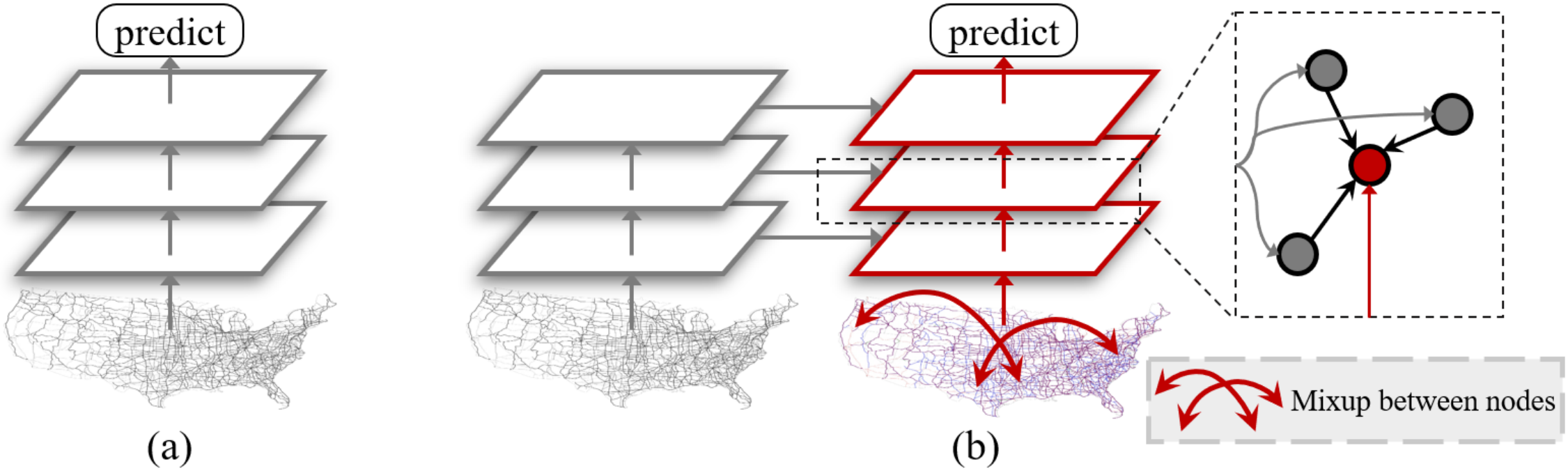
First contribution: Two-branch Mixup Convolution:



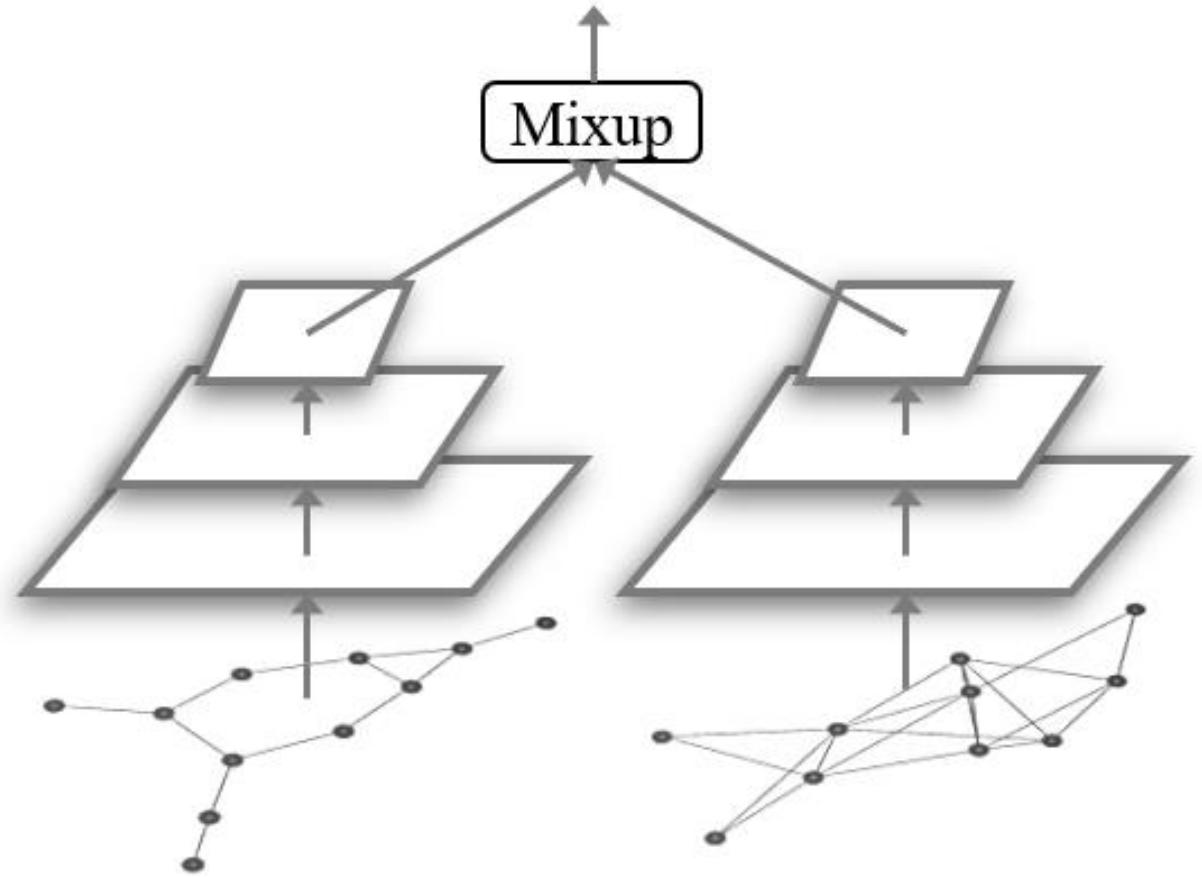
Conflicts between Mixup of Different Node Pairs due to Graphs' Connectivity



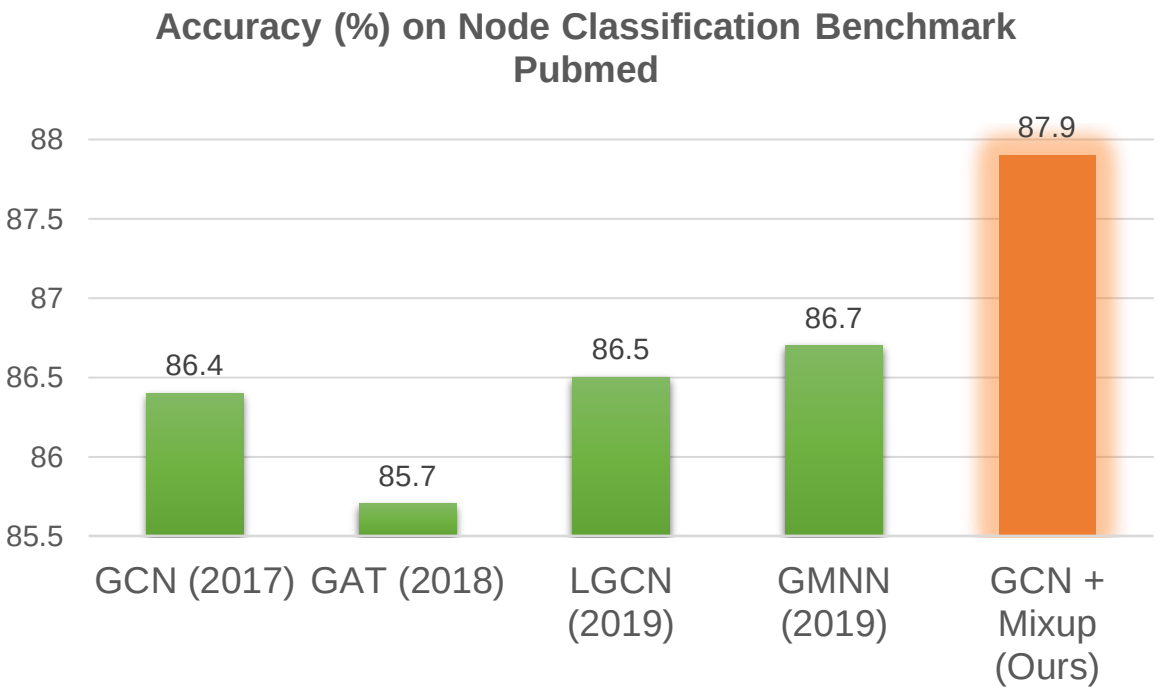
Two-stage Mixup Resolves Conflicts

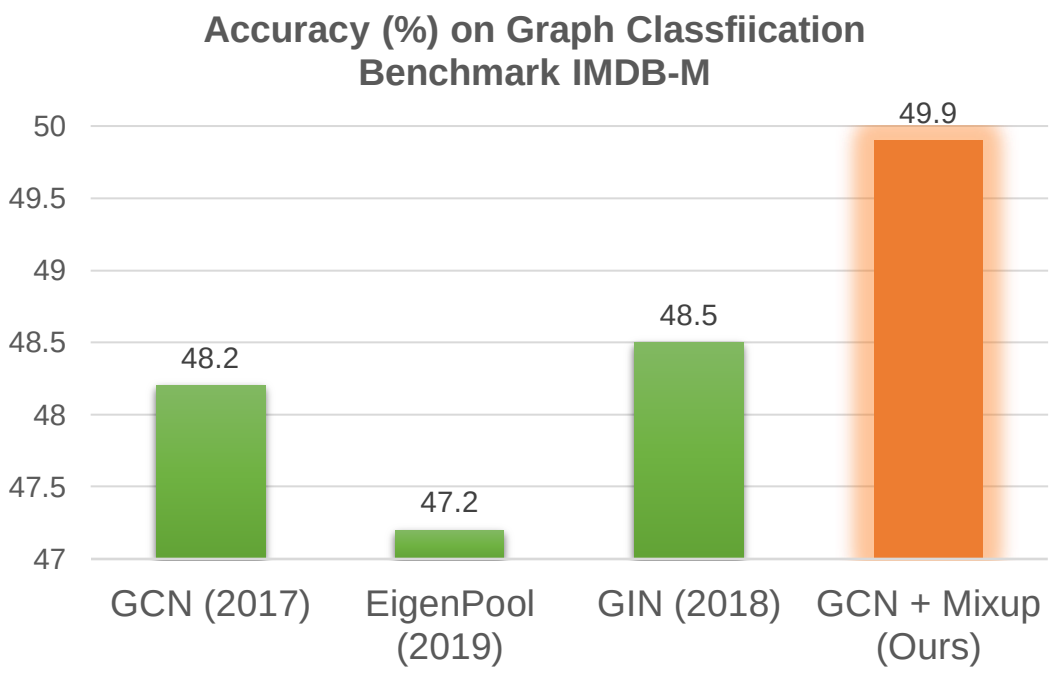


Mixup for Graph Classification



Experiments on Node and Graph Classification





GENN: Graph Explicit Neural Networks: Explicitly Encoding Graphs for Efficient and Accurate Inference (2023)

Graph Explicit Neural Networks:
Explicitly Encoding Graphs for Efficient and Accurate Inference

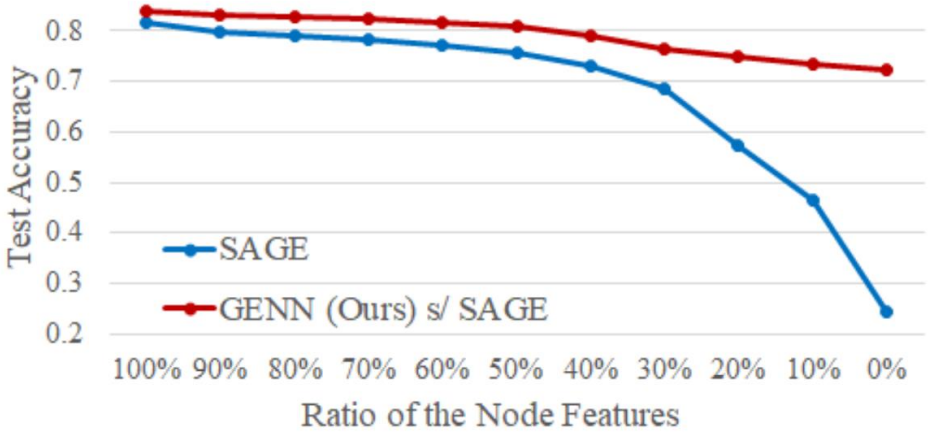
Yiwei Wang
National University of Singapore
Singapore
wangyw_seu@foxmail.com

Bryan Hooi
National University of Singapore
Singapore
bhooi@comp.nus.edu.sg

1. GENN, at the first time, resolves the serious performance degradation issues of GNNs on learning from the non-attributed graph learning.

2. GENN analyzes the poor performance of prior GNNs when learning from non-attributed graphs.

3. GENN is inductive and can provide efficient and accurate inference even on the graphs without node-level attributes.



LPNL: Scalable Link Prediction with Large Language Models (2024)

LPNL: Scalable Link Prediction with Large Language Models

Baolong Bi[1,3] **Shenghua Liu**[1,3*] **Yiwei Wang**[2] **Lingrui Mei**[1,3] **Xueqi Cheng**[1,3]

1. LPNL is the first paper that does the large-scale link prediction with LLMs.

2. LPNL describes graphs in the pure language format and resolves the efficiency issues when using LLMs to learn from graphs with a divide-and-conquer algorithm.

3. LPNL beats many GNNs on the large-scale link prediction.

Author Disambiguation Example

prefix_question: Which following candidate author writes the paper p_1?

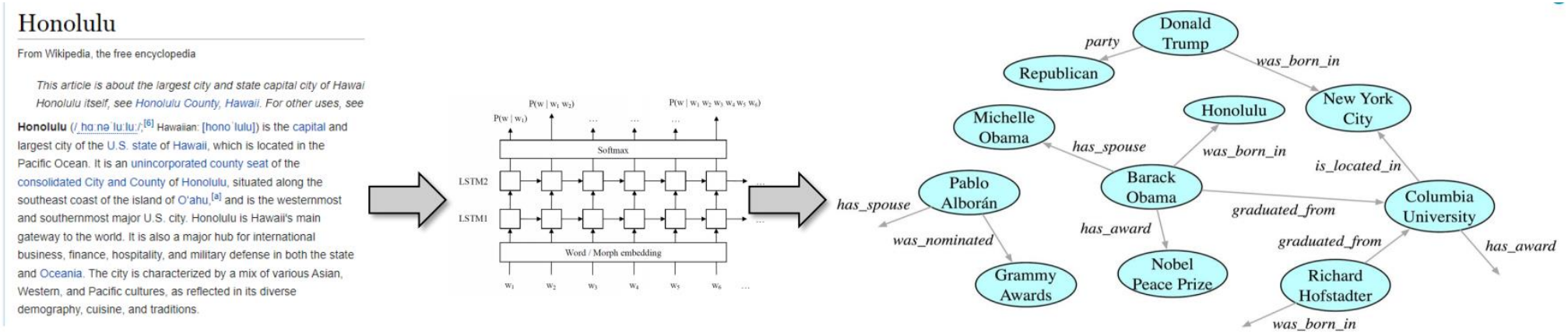
source_node_description: p_1: <*paper title*> is related with f_{25}: <*field name*>, v_{13}: <*journal info*>, p_{46}: <*paper title*>, a_{38}: <*author info*>, p_{27}: <*paper title*>...

candidate_nodes_description: a_1: <*author info*> is related with p_{15}: <*paper title*>...; a_2: ...; a_3: ...

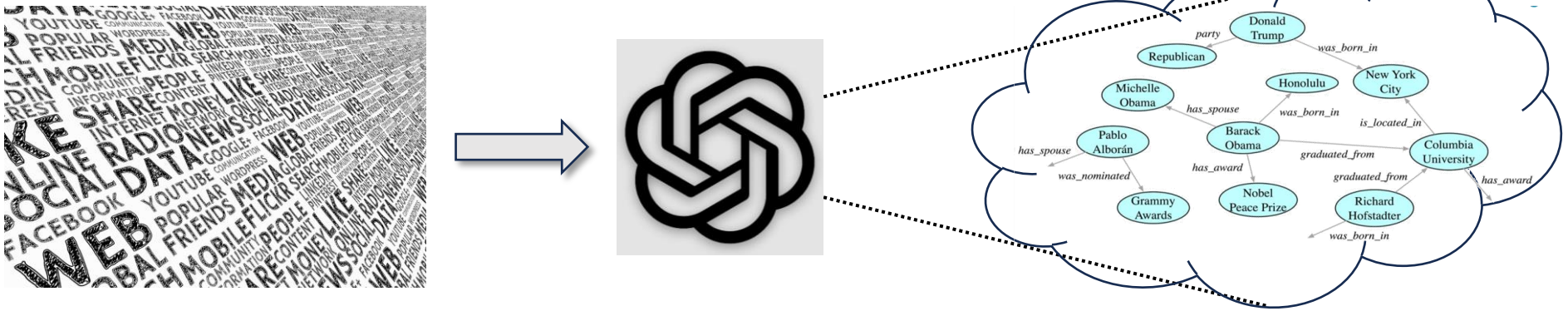
What' Next for Graph Learning? My Vision: Embracing Implicit Graphs

Researchers should not only insist on the "explicit" graphs, but embrace the "implicit" graphs that are stored as the parametric knowledge in the Large Language Models. Two kind of graphs can work together to help AI to understand relations.

Classical graph learning of "explicit" graphs.



Expected "implicit" graph learning.



DeepEdit: Knowledge Editing as Decoding with Constraints (2024)

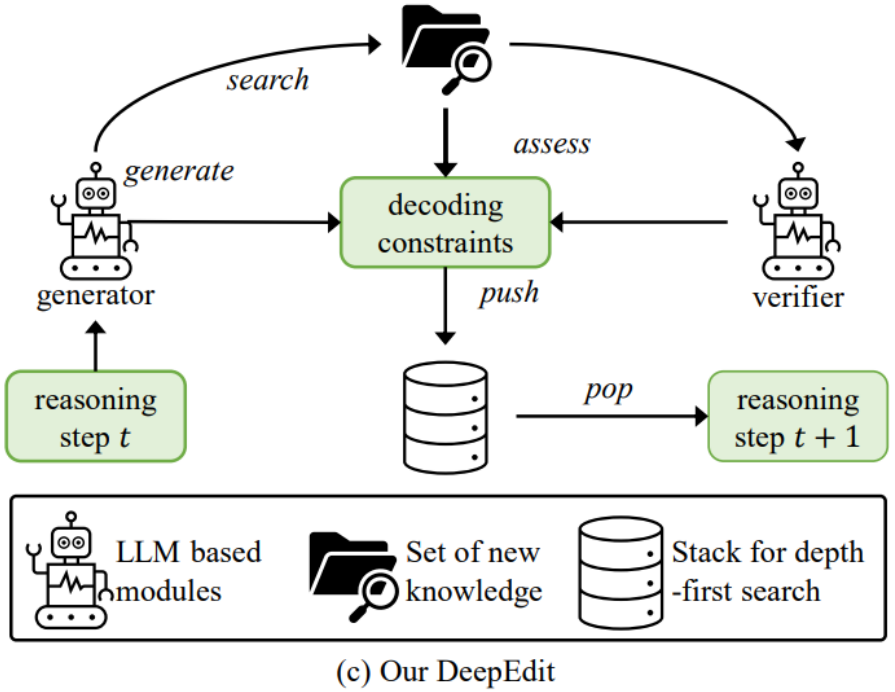
DeepEdit: Knowledge Editing as Decoding with Constraints

Yiwei Wang[†] **Muhao Chen**[‡] **Nanyun Peng**[†] **Kai-Wei Chang**[†]
[†] University of California, Los Angeles [‡] University of California, Davis
wangyw.evan@gmail.com
https://wangywust.github.io/deepedit.io/

1. Given the implicit graphs memorized by LLMs, we should be able to edit it like how we do for the explicit graphs.

2. DeepEdit edits the knowledge of LLMs with the updated real-world knowledge.

3. DeepEdit makes a step toward the effective editing of factual knowledge stored inside the implicit graphs memorized by LLMs.



(c) Our DeepEdit

Overview and Takeaways

Reviewing GNNs' Developments from 2016 to 2023:

Graphs are Young and Advanced Data to Describe Relations of Concepts

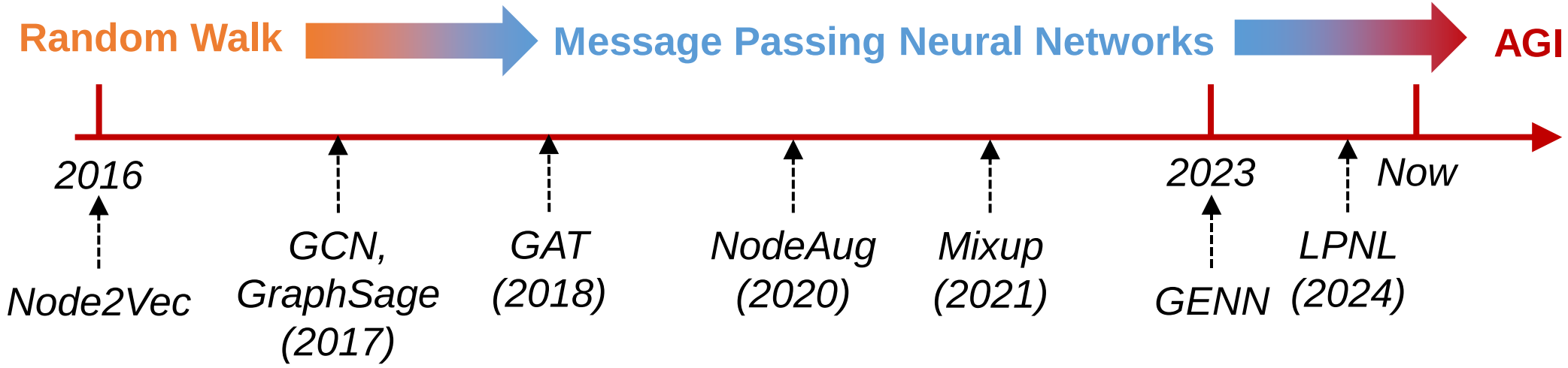
*The capability of describing connections of GNNs is the basis of its high performance on semi-supervised learning. The "**expert-made**" property makes graph data limited in nature.*

From 2016 to 2023, From Random Walk to Message Passing Neural Networks

Improved Learning Capability on Attributed Graphs.

From 2023 to future, From Message Passing Neural Networks to Artificial General Intelligence (AGI)

Instruction-Tuned LLMs Provide A Unified Protocol to Learn from All Data Including Graphs. Researchers may transfer the research focus from explicit graphs to implicit graphs.



Thank You!